Next generation text-mining applied to toxicogenomics data analysis = Next-generation text-mining toegepast op toxicogenomics data analyse

Stellingen

Behorende bij het proefschrift

Next-generation text mining applied to toxicogenomics data analysis

1. Knowledge-based filtering and disambiguation rules have a dramatic positive effect on the precision of dictionary-based biomedical and chemical term identification in scientific literature.
2. A dictionary of chemical terms is not enough to identify all chemical terms in scientific literature.
3. Gene sets based on concept profile matching provide a broader scope than gene sets based on manually curated databases.
4. There is not one perfect gene set analysis tool that fits every data set.
5. Until bioinformaticians are forced to accompany their journal publications with a preserved and executable implementation of their experiment, the discipline will continue to produce methods that almost no biologist will use.
6. A bioinformatics toolbox that does not include text mining is incomplete.
7. Publishing a bioinformatics method in a high impact journal without an accompanied new, exciting, and experimentally confirmed biological finding is practically impossible, and something that is holding back scientific progress in the field.
8. Data mining can be defined according to the principle of "my data is mine, and your data is mine". *With reference to C. Goble, The Seven Deadly Sins of Bioinformatics, 2007.*
9. Continuous application of pressure on the gas pedal will inevitably cause a crash.

Kristina Hettne